Journal of Universal Computer Science, vol. 21, no. 1 (2015), 156-174
submitted: 16/5/14, accepted: 1/10/14, appeared: 1/1/15 © J.UCS

Learning Analytics for English Language Teaching

Hannah Volk
(University of Graz, Austria
hannah.volk@uni-graz.at)

Karl Kellner
(Private University of Teacher Education KPH Graz, Graz, Austria
karl.kellner@kphgraz.at)

David Wohlhart
(Private University of Teacher Education KPH Graz, Graz, Austria
david.wohlhart@kphgraz.at)

Abstract: In recent times, online learning platforms get more and more attention and the number of collected data is growing. Learning analytics is a valuable opportunity to gain specific information for a better understanding of student's learning behaviour and to improve their learning success. In this work, the collected data of the online learning platform *www.more-online.at* is analysed and first research results are presented. The main objective is to analyse the usage behaviour over a school year and to show the diffusion of the online platform among provinces in Austria, different school types and other characteristics. Furthermore, the content of the online platform is put under closer examination to enable decisions about the efficiency and effectiveness of different types of exercises.

Keywords: Learning Analytics, Big Data Analysis, Educational data mining, E-learning
Categories: L.3.5 - Online Education, L.3.0 - eLearning Systems/Technology/Tools/Platforms

1 Introduction

The technological revolution in the late 90s led us to a new dimension in the field of education. The advancement of technology fosters the demand of online learning. [Kakoty et al. 2011] Nowadays learning more often happens via the Internet. In schools, learning platforms are increasingly regarded as a useful and necessary addition to school education. Publishers of schoolbooks and other providers constantly extend their offer. The user frequency is growing. This rapid rise of e-learning during the last years also led to an increase of available user data. Educational institutions and teachers can take advantage of this development and get the chance to use these data as a basis for their decisions. The research area of learning analytics aims at a development of methods and techniques to analyse these data. [Neuhold 2013] Based on the fact that information technology strongly influences the development of human society, it is indispensable to take advantage of modern technological facilities in order to support and improve English language education. Students learning English as a second language require language support in the competences hearing, speaking, reading and writing in order to gain experience and improve their skills. [CEFR 2011] The use of various tools enables an easier and

more effective process of learning the English language. [Nomass 2013] The importance of technology in the context of learning a language, as an addition to the role of teachers, is growing. This combination can bring about advanced learning results of English learners. [Sharma 2009] These positive effects of learning with technology on student outcomes were demonstrated in different meta-analyses, such as those conducted by [Waxman et al. 2003]. The results of their study about the effectiveness of technology on student's learning process show a modest, positive effect of teaching and learning with technology on student outcomes. [Waxman et al. 2003]

English language education has dealt with the use of technology for some time in recent history. Many educational settings in the sixties and seventies made use of language laboratories where teachers were able to monitor their student's conversations. Although the language laboratory was a positive step in connecting language education and technology, it was soon considered as laborious and boring for learners. These aspects caused changes in the communicative approach to second language education, and computer assisted language learning (CALL) was introduced. CALL software provided a new medium for language learning and the educational literature widely covers its advantages as a teaching and learning tool, due to the interactive and multimedia capabilities of the computer. Nowadays CALL software is designed to include sound, graphics, video, and animation, which makes it an attractive tool for teaching and learning. [Shingal 1997]

In this article, we discuss the application of Learning Analytics in the field of English language teaching and present first results from the field. The following work is divided into five parts, starting with an introduction [Section 1] including general information about the research area and providing context information of our study. In [Section 2] the main research questions are presented. [Section 3] and [Section 4] contain our main results on the usage of the online platform and the learning behaviour of English learners. Some conclusions concerning the current work are made in [Section 5] and future work is discussed in [Section 6].

2 Focus of this Study

In the present study, the online learning platform *www.more-online.at* for English learning is analysed for the purposes of understanding and optimising the learning behaviour of students. Up to 20.000 students per day use this platform, doing their English-homework. On average 150.000 users per month visit www.more-online.at. The platform is an additional tool to the course book *More!*, which is used in Austrian schools for English education at lower secondary education level. [Gerngross et al. 2008] Since the availability of the schoolbook More! is limited to Austria, the analysis will focus on the Austrian school system. The focus of this study was put on the book More! 1, which is allocated to students of the first year of lower secondary education. The online course encompasses 40 so called *Cyber Homework* units including 159 exercises. A Cyber Homework unit is a set of online, interactive homework exercises, which enable students to do their homework online. Each Cyber Homework unit consists of a set of interactive exercises, which cover specific competence areas such as reading, writing and speaking and are to be completed in some 15 to 20 minutes. This period has been chosen in accordance with findings that

20 to 40 minutes of homework time is recommended in grades 4 to 6. [Brewster and Fager 2000] In order to enable students to use Cyber Homeworks, teachers have to define classes, hand out the credentials to each student and allocate a certain Cyber Homework course to the class. Then teachers assign a Cyber Homework unit to a specific class by setting a deadline for the completion of the homework. When logging in, students get an overview of homework that is assigned to them together with the submission deadlines. Teachers see at a glance which students already submitted their homework, and get the assessment results from the online platform. There is an overall score for each exercise, which is assessed automatically on the platform once the homework is completed. All results are shown online for a particular class so teachers can gain an overview of a class's work and provide feedback via the integrated messaging system or assign extra practice. Students can redo an exercise until the deadline set by the teacher expires. Once the deadline expired, students get a feedback about their results and they are provided with detailed information about the correct answers and their mistakes. This design was chosen because informing students about their learning process is one major requirement in order to increase their motivation. A successful learning process requires feedback about the quality of ones learning performance. [Grasedieck 2008] Feedback, which is more detailed and goes beyond a simple feedback of results, reinforces the positive effect on learners' motivation. Learners thereby get informed about the effects of their learning performance and regard their efforts as noticed and appreciated, which leads to an increased feeling of competence and joy of learning. [Möller and Wild 2009]

The analysed data include user data of the online platform starting from June 2008 to August 2013. Regarding the development of usage, a strong increase of distinct general users and new users is visible over the years, which is due to the extension of the online platform's offer. Therefore, the focus of this study is put on the school year 2012/13 (September 2012 to June 2013]) where the full offer of exercises was available on the online platform. The main research questions of this prevailing study are concerned with two different subject areas, the learning process and learning outcome of students. First, the user behaviour at the online platform is of major interest, in order to obtain a general overview of the online platform and to track the behaviour of website visitors. Our target is to examine how students engage with and study in the e-learning environment. For the purpose of taking an inner and outer perspective on the usage behaviour, the number of visitors, as recorded by Google Analytics was compared with the number of actual exercise results produced on the online platform. The second subject area of this study is a detailed analysis of the exercise results produced, in order to enable decisions on matters such as efficiency and effectiveness of different exercises. Therefore, the apparent level of difficulty of the different exercises in More! 1 is evaluated, by analysing their average success rates. Another interesting aspect in relation to the learning outcome of students is a closer examination of student's learning behaviour by analysing the repetitions of exercise attempts and improvements of exercise results. The following chapters of this article will enable a closer insight into the realization of these research questions and outline its major results.

3 User Behaviour from a Global Perspective

Throughout this section, we present an overview of user behaviour on the online platform *more-online.at.* First, the factor time is put under closer examination and the development of user behaviour within a year, week and day is analysed. System based records of users activities are automatically captured by e-learning environments, giving information about who accessed what, and when. [Phillips et al. 2011] This information can be of high importance for publisher of schoolbooks and developers of e-learning programmes as much as for teachers. The results can contribute to improvements of the online platform's content and usability and to an adaption of the content according to student's needs. Improvements of the course quality enable students the use of actualized and optimal educational material and, therefore, higher performance in exams. [Karakos and Kazanidis 2011] As reported by [Phillips et al. 2011] these 'usage logs' of students using E-learning applications have been under-utilised in E-learning research". The study of this automatically captured data, which records who accessed what and when to study student behaviour, is now termed learning analytics." [Phillips et al. 2012, 2861] The techniques of learning analytics enable direct access to informations on student's learning behaviour, which is an advantage to more qualitative approaches (e.g., interviews) where the information is filtered by the perceptions of the student. [Phillips et al. 2012] According to [Judd and Kennedy 2004], electronic records from technology based learning environments enable researchers and evaluators to recognize different patterns of usage and they present empirical evidence to prove this by giving explanations about how usage data can be processed and analysed in an effective way.

In their case study [Karakos and Kazanidis 2011] deal with an Open eClass e-learning platform used by Greek university students and describe the use of data mining techniques to enable an analysis of usage logs and draw useful conclusions. They point out the great value of these techniques and explain how an analysis of usage logs contributes to improvements in the content and usability of e-learning environments. The utility of information derived from usage logs was further recognized by [Ben Naim et al. 2009], who conducted research about data mining techniques and data visualization in an educational e-learning environment. They developed a tool, which is able to visualize student's behaviour and provides visual feedback about the effectiveness of adaptive learning elements.

In their recent study [Kaur and Krishan 2013] conducted an analysis of student's behaviour by making use of data mining techniques and tools. In order to analyze the relationship between how students use a course and how they perform at school, the researchers used various clustering and classification procedures. [Kaur and Krishan 2013] [Shen et al. 2002] who developed a Data Analysis Centre based on an e-learning platform chose another interesting approach. Their tools support teachers in the analysing process of student's learning behaviour and process and enable an efficient organization of its web-based contents.

In their pilot study about learning analytics and study behaviour, [Philips et al. 2011] investigate the engagement of students in an e-learning environment. They point out that usage logs however do not give us explanations about the reasons for user's behaviour; they just record the activities and behavioural responses of users. Analysing and interpreting this data should therefore be carefully done and results

should be critically scrutinised on the basis of various touchstones. [Phillips et al. 2011]

3.1 Temporal Aspects of User Behaviour

In order to obtain a general overview of the online platform and to track the behaviour of website visitors the tool Google Analytics is used. Google Analytics records access data to any page of the course content. Further, these results are compared with the actual exercise results produced by users of the online platform. This enables a comparison of an inner and outer perspective on the platform. First, it is interesting to look at the use of the platform within a school year. [Fig. 1] shows the amount of distinct visitors at the online platform, where each visitor is only counted once (blue graph) as much as the amount of exercise results produced (green graph) in the year 2012. In general, the line graphs show a quite similar behaviour and illustrate the peeks and deeps throughout a school year caused by holidays and vacation time. A school year in Austria is divided into two periods, a winter and a summer term. The winter term starts in the beginning of September and ends in February (the actual date depends on the province). After a semester break of one week, the summer term starts and goes along until July, when the summer holidays begin.

The description of [Fig. 1] starts in September (week 36), in order to give a better picture of a school year. The beginning of a new school year in September (week 36) leads to a steady increase of usage (in terms of exercise results produced) until the peak is reached in the middle of October. Then the curve shows a small decline, probably caused by the autumn holidays taking place in some areas. Afterwards the usage rises again and stays on a high level until a sharp decrease during Christmas holidays (week 51). After a consistent level of usage during January, the winter break in the beginning of February (week 5-7) leads to a visible decline. A much stronger decline in user's activity exists in March, ending up in a deep duri ng the Easter holidays (week 13). The rest of summer term shows a consistent level of usage with a last peak in the end of May until the usage finally continually declines the closer it comes to the end of the school year. During summer holidays, the green line graph reaches the neutral point demonstrating that virtually no exercise results are produced during the summer holidays. The blue line graph shows some small activity of users over summertime caused by search engines or visitors entering the online platform without performing exercises.

In order to find explanations for the decreasing user activity at the end of a school year, it is interesting to investigate the frequency of usage of the different exercises in More! 1. This aspect is relevant in order to find out if the decreasing user activity can be explained by the fact that schools are simply done with all Cyber Homeworks before the end of the semester. The results of the regarded sample make clear that the exercises, which are used least frequently, are those belonging to Cyber Homeworks that are located further back in the course book More! 1 and are intended for the last weeks of the school year. Therefore, the steady decrease of usage at the end of a school year cannot be explained by the completion of all Cyber Homeworks but may be a general phenomenon at this point of time.

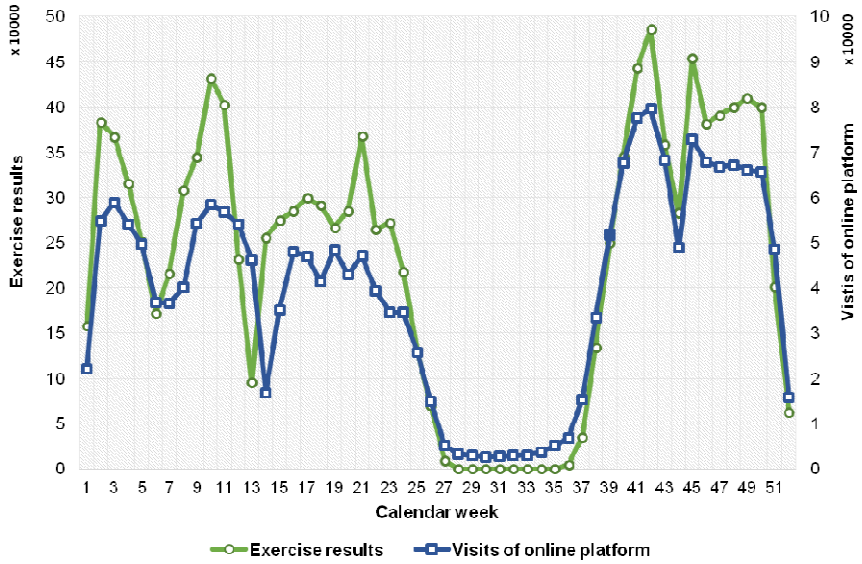

Figure 1: Usage of the online platform during the year 2012

In the following step, the distribution of usage throughout a general school week is analysed. Inner and outer perspective show quite similar results, whereas the distribution of visits of the online platform aligns with the number of exercise results produced. The results in [Fig. 2] indicate that the usage is on a constantly high level from Monday to Wednesday. In the second half of the week, starting on Thursday, a gradual decline can be noticed and the minimum is reached on Saturday. Finally, the usage distinctly rises again on Sunday, reaching almost the same level of usage as during the beginning of the week. The activity on the weekend therefore strongly differs. Whereas Saturday shows only little user activity on the online platform, Sunday seems to be the day when many students do their homework or prepare for exams by using the online platform.

A possible explanation for the more intense usage on Sundays could be found in the individual homework deadlines set by teachers. A closer analysis of the user behaviour indicates that Sundays are most often used as final days to deliver homework on the online platform. As indicated in [Tab. 2], 21.88% of all deadlines in the school year 2012/13 were Sundays. In contrast to that, only 5.70% of all Cyber Homework units were required to be done by Saturday. During the weekdays, the amount of deadlines was on a consistent level, only on Mondays it was a bit higher than during the rest of the week.

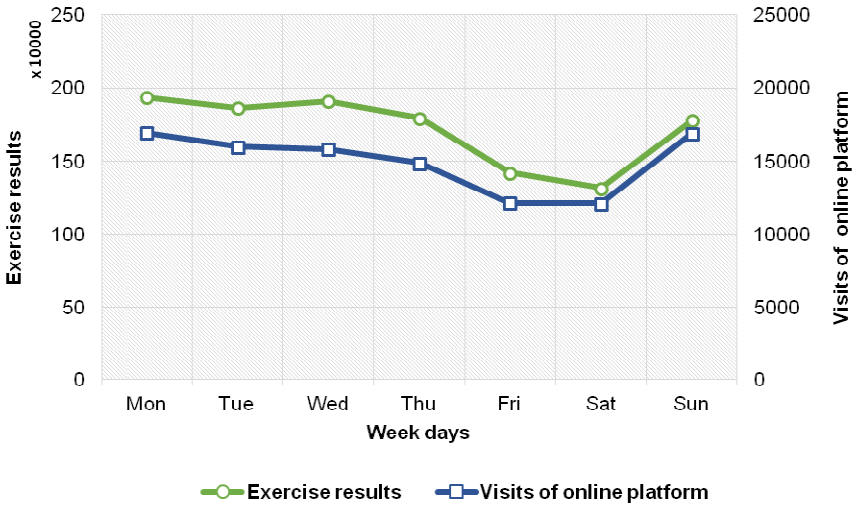

Figure 2: Usage of the online platform during a general school week

	Frequency	Valid Percent
Monday	9612	16,19%
Tuesday	8229	13,86%
Wednesday	7818	13,17%
Thursday	8667	14,60%
Friday	8678	14,61%
Saturday	3383	5,70%
Sunday	12993	21,88%
Total	59380	100,00%

Table 2: Deadlines for Cyber Homeworks of More 1 in school year 2012/13

Finally, a closer look shall be taken at the usage behaviour during a general school day. As follows from [Fig. 3] students start their activity at around 6 am. The usage shows a sharp rise during the morning until a first peak is reached at 10 am. From 10 am to 12 am a slight decline in usage is visible. When looking at the visits the distribution leads to the assumption that use in schools is only a small part of the overall activity. Looking at the inner perspective (exercise results produced) makes clear that the actual usage activity in the morning at school is considerably higher than expected from the outer perspective. The difference can be explained by the way Google Analytics counts access from (school) networks. Since these networks use a common IP address, they are considered as a single distinct user, so they are counted only once. Regarding historical developments, first computer programmes supporting student's learning process were primarily produced for independent learning and

practising at home without an instructor. Communicative, action-oriented and cooperative learning forms made an application of learning software in class seemingly difficult and less valuable. With the rise of the Internet and different application programs with a more instrumental than tutorial character, this conception started to change. [Nandorf and Schmidt 2003]

Nowadays recent studies show, that the usage of computers in classroom is growing. The KIM-Study 2012, a general study about the use of media among 6- to 13-year-old children in Germany indicates, that besides an increasing usage of computers at home for school, computer and internet are also used during class at school more frequently. All together 44% of 6- to 13-year-old children already gained experiences with a computer at school. The use of computers in class rises with children's age. As far as the computer usage at school is concerned the writing of words and texts is the most common activity. On the second place is the use of learning software, which is regularly used by 2/3 of the interviewees. However, it is necessary to emphasise that the KIM-Study investigates independent learning with learning software. [Medien- pädagogischer Forschungsverbund Südwest 2013]

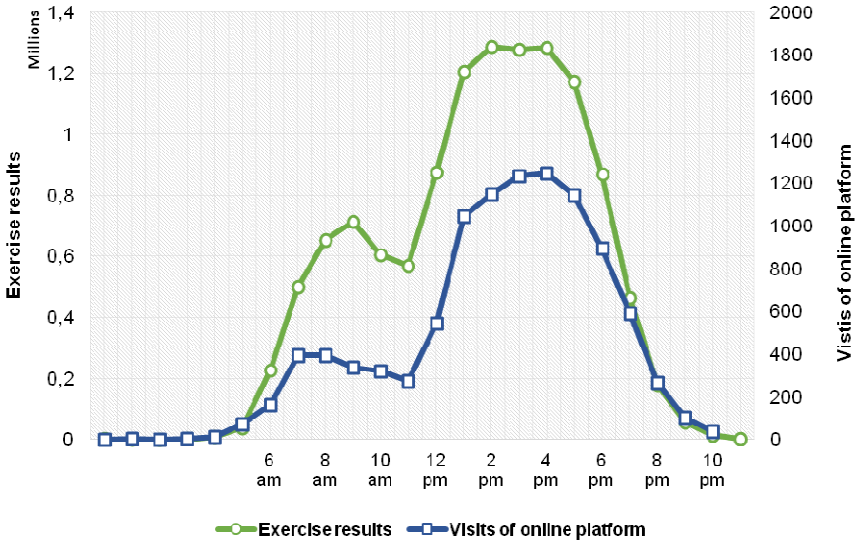

Figure 3: Usage of the online platform during a general school day

As indicated in [Fig. 3] above, the green graph line demonstrates another strong rise of the amount of produced exercise results from 12:00 pm on. The peak of the day is reached between 3:00 and 5:00 pm. This is probably the time when most of the students do their homework and practice for exams. From 5:00 pm on, the usage activity is under a steady decline reaching the neutral point at around 10:00 pm.

3.2 Educational Aspects of User Behaviour

The following section aims at a description of platform's users. In order to enable a clearer understanding of the following results, some context information about the

educational system of Austria is necessary. In the Austrian school system, compulsory schooling starts in September, following a child's sixth birthday and lasts nine school years. The education for children is divided into three main categories: primary, lower secondary and upper secondary. Primary education (Volksschule) lasts four years. At lower secondary level, which lasts four years, children can chose between three further school types: *Allgemeinbildende Höhere Schule* (AHS), *Hauptschule* (HS) and *Neue Mittelschule* (NMS). The school types HS and NMS are quadrennial and afterwards students either go to a Polytechnical School and learn a profession or they continue their education in the upper cycle of an AHS. Upper secondary education lasts four to five years and refers to AHS upper cycle (4 years) and vocational secondary education (five years). All streams lead to the Matura, which gives access to higher education. [BMUKK 2014] Since school year 2008/09, one major education policy measure in Austria is the transformation of HS into NMS in order to create a common school for all 10- to 14-year-old children. The NMS tries to realize modern educational concepts and to design a new common culture of learning. The transformation process should be finished until 2015/16, when all HS should have developed into NMS. [BMUKK 2013]

The focus is now put on the distribution of usage throughout different school types in Austria. Since More! 1 is allocated to the lower secondary level, it is mainly used by the school types Allgemeinbildende Höhere Schule (AHS), Hauptschule (HS) and Neue Mittelschule (NMS. Because the transformation of HS into NMS continues to move forward and many transformations happened during the regarded school year 2012/13, these two school types are now put together. The results show that schools of the school type AHS seem to make greater use of the online platform than those of HS/NMS do. Overall, more than three quarters of all AHS in Austria (82.67%) used *more-online.at* during the school year 2012/13. Regarding HS/NMS, about half of the schools (50.85%) of these school types made use of the online platform in this period. [Statistik Austria 2013]

3.3 Topographical Aspects of User Behaviour

The following paragraph deals with the distribution of usage among the provinces of Austria. Austria encompasses nine provinces: Lower Austria (LA), Upper Austria (UA), Styria (ST), Vienna (VIE), Tyrol (TY), Salzburg (SL), Carinthia (CA), Vorarlberg (VO) and Burgenland (BL). Since each school in Austria has a certain indicator number it is possible to figure out their location in Austria. In total 1062 schools used the online platform during the school year 2012/13, whereby usage is defined as the performance of at least 200 exercise results. The vast majority of them (1036) are schools at lower secondary level (AHS, HS/NMS). Therefore 55.57% of all Austrian schools in this sector used the online platform during the school year 2012/13. [Statistik Austria 2013] The columns in [Fig. 4] demonstrate the percentage of schools in each province using the online platform during the school year 2012/13 in relation to the total number of schools in each province. The data used derives from the educational statistic of 2012/13 by Statistik Austria. [Statistik Austria 2013] Those provinces showing the highest percentage of schools using the online platform are Burgenland (BL), Salzburg (SL) and Vorarlberg (VO). The three provinces with a lower percentage are Upper Austria (UA), Vienna (VIE) and Tyrol (TR).

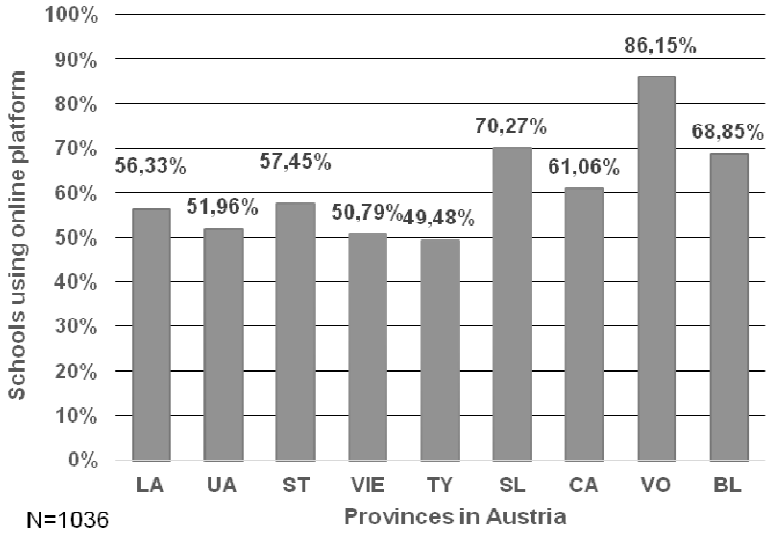

Figure 4: Distribution of usage among Austrian province (school year 2012/13)

3.4 Explanatory Model of Usage

Based on the results of the first data analysis described above it is now interesting to figure out which factors influence the activity of a school at the online platform. The activity of a school is defined by the number of performed exercises divided by the number of users, who produced at least one exercise. In order to answer this question a binary logistic regression model is calculated. Through a logistic regression it is possible to determine the impact of multiple independent variables to enable a prediction of the membership of the dependent variable categories. The logistic regression aims at finding the best fitting function using the maximum likelihood method, which maximizes the probability of classifying the observed data into the appropriate category. [Backhaus et al. 2011] As dependent variable, the activity of a school is defined, calculated as described above. The variable is dichotomized in order to differentiate between schools with a small and large number of exercise results per user, whereas the average amount of learning results per student was used as distribution criterion. As independent variables the school type, province, region and school provider are used. [Tab. 2] demonstrates the final model with all independent variables showing a significant influence on the activity of a school.

According to the results, the general number of users of More! 1 at one school is a major influence factor of the activity of one user. This means that a user is more likely to have a large amount of exercise results if the number of users per school is higher. This result suggests that the use of the platform is dependent on cooperative effects, possibly between students or between teachers. It suggests also that use of media can be fostered by school-wide adoption strategies. In addition to that, the school type is another factor of influence on the results. Schools that belong to the type Allgemeinbildende Höhere Schule (AHS) thereby show a larger amount of

exercise results per user than other school types. According to the findings of [Bitkom 2011], a representative study on the use of electronic media at schools in Germany, teachers of the school type AHS have a more positive attitude towards the use of electronic media at school than their colleagues in other school types do. Those teachers with a more negative attitude towards the use of electronic media are primarily teachers that are more than fifty years old. These so-called Digital Immigrants got to know computers, internet and other digital technologies not until adult age. [Bitkom 2011]

Coefficients					
	Unstandardized Coefficients			Standardized Coefficients	
Model	**B**	**Std. Error**	**Wald**	**Beta**	**Sig.**
Users of More 1!	.022	.003	46,387	1.022	.000
School type AHS	1.101	.220	25.141	3.009	.000
Province Lower Austria	.685	.184	13.895	1.983	.000
City school	-.468	.167	7.880	.626	.005

Table 2: Logistic regression model

Another important factor is the province in Austria the school belongs to. Schools, which are located in lower Austria, are more likely to have a larger amount of exercise results per user than schools in other parts of Austria. Lower Austria is Austria's biggest province, it is located in the north and it encompasses Vienna. This province has a special location as it actually lacks a real capital city because the biggest urban conglomeration in this area is Vienna and is therefore located outside Lower Austria. Another possible explanation could be the fact that Lower Austria is one of those provinces with the highest density of computers in households. Thus, 79.9% of all households in Lower Austria hold at least one computer, what is slightly more than in most of the other provinces. [Statistik Austria 2011]

Finally, the analysis made clear that the location of schools has an influence on user's activity. Thus, schools at the countryside are more likely to show a higher activity per user than schools located in urban areas like cities and towns. A possible explanation for this result might be found in the OECD report 2006 about information and computer technology. The study covers the extent of ICT resources at school in 30 countries worldwide and focuses, among others, on the question, how much the ICT resources depend on the school's location (rural locations and towns versus cities). The results show that in most countries no differences in the number of computers per student between schools in rural locations or towns and schools in cities exist. However, in Austria and some other countries a tendency is visible that schools in rural locations or towns have more computers per student than city schools. [OECD 2005] Another interesting aspect is the availability of computers and internet access in Austrian households. According to a study of Statistik Austria in 2011, there is no significant difference between towns/cities or rural areas concerning computer equipment. However, rural locations show a lower number of households with

internet access than towns or cities, although these differences are lower than one would expect. [Statistik Austria 2011]

In order to evaluate the goodness of the final model, some specific diagnostic tests are performed. The results of the Hosmer and Lemeshow Test for goodness of fit (p=.369, ns.) indicate that the model fits the data well. The classification table also shows an appropriate result. After an optimization of the cut-off-value, 65.6% of all occurrences are correctly predicted. The Nagelkerke R squared indicates that 20.2% of the variation in the dependent variable is explained by the model, what can be classified as quite satisfactory. [Backhaus et al. 2011] Nevertheless the explained variation is not as high as expected beforehand. A large amount of variance therefore has to be put down to other factors or groups of factors, which are not available in this study.

4 Learning Outcome

The following chapter contains a closer examination of the exercise results, which were produced by students doing More! 1 Cyber Homeworks. In order to do so, a sample of four schools was selected. The number of exercise results produced by each school was determinative for the choice. The four schools with the most exercise results for More! 1 were included. As indicated in [Tab. 3], the regarded sample encompasses 582 students, who produced 74.872 exercise results in More! 1 in the school year 2012/13.

	Exercise results More 1	Number of users
School 1	21691	168
School 2	24051	179
School 3	15556	114
School 4	13574	121
Total	74872	582

Table 3: Regarded sample of schools (school year 2012/13)

The first interesting aspect of this analysis is to get information about the level of difficulty of the different exercises in More! 1 by an examination of their average success rates. The average success rate includes all attempts by all students of the regarded sample. As indicated in [Fig. 5], the largest proportion of exercises in More! 1 (39.62%) leads to an average success rate between 90-100 percent. The second largest amount of exercises shows average success rates between 80-89 percent. According to that the vast majority of exercises (almost three quarters) in More! 1 show comparatively high average success rates. In contrast to that, only about one quarter of the exercises leads to average success rates of less than 80 percent.

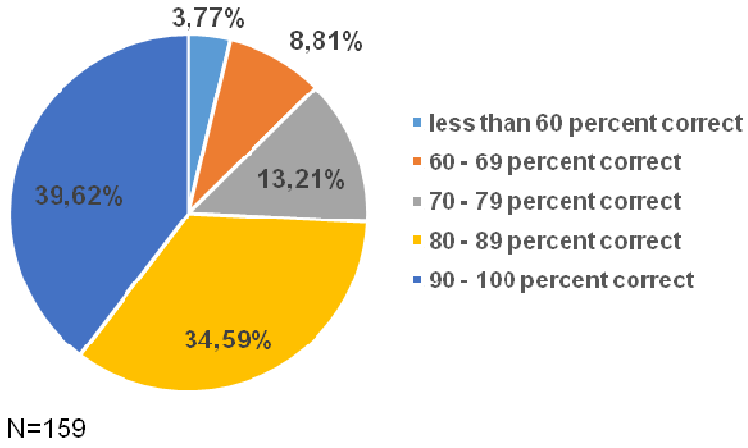

Figure 5: Average success rates of the exercises in More 1

In order to enable a better classification of the exercises, a scattering of values around an average success rate between 75 and 85 percent was defined beforehand as the optimal range. This is due to the assumption that success rates in this area lead to the highest motivation and thus to the highest learning ability of the students. [Keller and Suzuki 2004] line out the importance of confidence and positive expectancies for success as a key factor of motivation. An average level of difficulty of the learning requirements has a positive influence on the learning success. Homeworks therefore should be neither too easy nor too difficult. During this part of the learning process, certain psychic forces are being mobilised, which are required in order to deal with the learning content. [Grasedieck 2008] According to [Meier 2006], who deals with the psychology of learning in connection to E-learning, students should neither be subchallenged by a monotonous repetition of what they already learned, nor overstrained by highly demanding exercises.

In the following step, the focus is now put on the learning behaviour of the students of the selected sample in order to figure out their learning outcome. The results indicate that on average about 16% of all exercise attempts performed by the regarded sample are repetitions. Therefore, students seem to repeat exercises when they are not pleased with their performance, although they are not informed about their achieved results right away. According to the correlation analysis performed, those exercises showing lower average success rates, are repeated more often than those where the average success rates are higher (r=-.594**). In addition to that, students show the higher improvements of their success rate, the lower the average success rate of an exercise is (r=-.439**). Thus, more difficult exercises enable a stronger improvement of exercise results. Furthermore, also the number of repetitions goes along with the improvement of exercise results, whereas those exercises repeated more often also show higher increases of success rates (r=.329**).

In this regard, it is interesting to take a closer look at the exercises results in the first and last attempt of each student in order to figure out the development of success rates. [Fig. 6] includes the success rates of the first attempt (blue bar) and those of the last attempt (green bar). The results show that the repetition of exercises leads to a strong improvement of their success rate. Thus, 32.14% of all first attempts lead to success rates between 90-100%. Regarding the final attempts, this value is more than twice as high (73.03%).

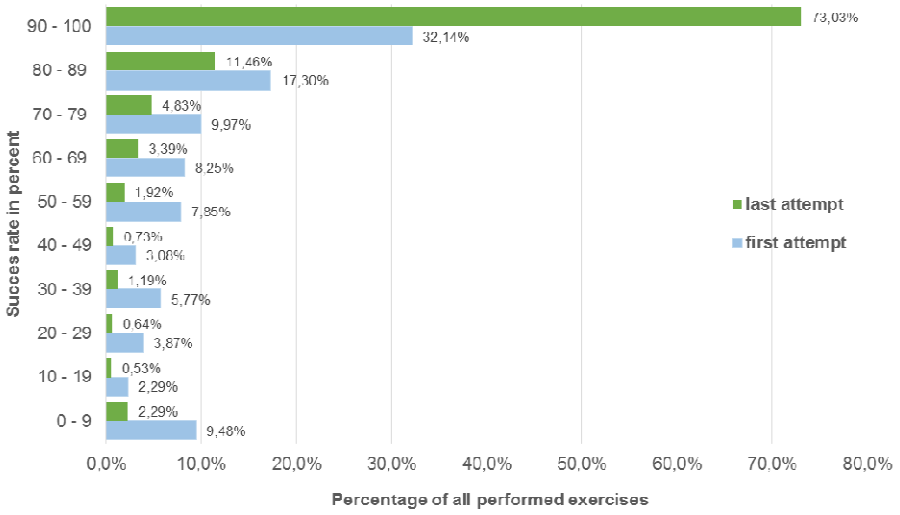

Figure 6: Success rates of first and last exercise attempts

A closer look at the exercises produces more details. Since all Cyber Homework exercises of More! 1 are self-evaluating by design, there are no open-ended tasks and a limited exercise typology is used. Among the 159 exercises of More! 1 there are 53 matching, 44 multiple choice, 41 gap text and 13 dialogue order exercises. The average success rate for multiple choice and matching exercises is 85% whereas the average of gap text and dialogue order exercises is significantly lower: 71%. Gap texts are more difficult because writing is always error prone. The experienced difficulty of exercises where parts of a dialogue have to be assembled suggests that combined cognitive and language competences are necessary to discover the logical structure of a communication whereas the handling of the exercise is very simple. The easiest exercise (99%) is a drop down multiple-choice exercise where the students have to make a choice between two names of animals, the most difficult exercise (42%) is a dialogue order exercise about watching TV. A comparison between a subjective estimation of the exercise difficulty and the exercise results shows a high overall correlation of $r=0.523^{**}$. A detailed look at the differences between the estimation and the results can be a valuable source for improving the quality of E-learning content. Here an exercise was discovered that was estimated as difficult but showed very positive results (95%). The exercise asks the student to find several

words in a letter grid. The examination of the discrepancy showed that the exercise could be solved by simply clicking all letters in the grid.

5 Conclusions

This paper presents first research results of analysing the learning behaviour of students in an E-learning environment in order to contribute to a better understanding and an optimisation of their learning process. The online learning platform, which is analysed in this paper, is called *www.more-online.at.* The main content are Cyber Homeworks, a series of exercise packages closely related to the course book More! which is designed for English teaching at lower secondary education level. The study has gone some way towards understanding the learning process and learning outcome of students interacting with the online platform. In order to obtain a general overview of the platform's usage and to track the behaviour of website visitors, the usage behaviour was analysed by regarding the platform's visitors, as recorded by Google Analytics as much as the actual exercise results produced on the online platform.

The results show that the usage behaviour in an e-learning environment is strongly influenced by the factor time and the time and activity structure of a school year. Students show fewer interactions with the platform in weeks before, during and after school holidays. During the long summer break the activity almost reaches point zero. Furthermore, we found out that the platform is used more frequently during the first half of a school week, whereas in the second half the use is significantly lower. On the weekend, only little user activity is visible on Saturday whereas Sunday seems to be the day when many students do their Cyber Homework at the online platform.

Our findings about the temporal aspects of user behaviour gave first insights into how students and teachers interact with the platform and provide a valuable basis for future research. The results constitute an important basis for the development of further analytical instruments in the field of Learning Analytics. In terms of a triangulation of methods, qualitative empirical research methods could further be used as an addition to the quantitative approach of this study. Usability tests or customer inquiries would for example allow further insights into specific usage scenarios and interaction procedures. The results of our analysis provide a valuable basis for the development of these further research instruments and processes. A more detailed information about the needs and requirements of users allows a more specific adjustment of the platform's content. In the long term, this leads to a higher satisfaction of students and teachers and thus to a stronger commitment to the online-platform, what is of course in the interest of software-developers and schoolbook authors. Furthermore, the information gained about the usage behaviour in the course of time gives some indication about suitable test periods where new features or advanced contents can be implemented as prototypes and tested by potential customers. As a recommendation to other researchers, who are able to intervene into the process of data generation, we would like to add that a consideration of time courses e.g. performance time is another interesting aspect. This information is not available at the online-platform more-online.at which we investigated, so we couldn't cover it in our study. We regard research aspects focusing on this aspect as promising but also want to call attention to a possible data bias due to differences in the working speed and attitude of learners.

Another major focus of this paper was an analysis of the distribution of e-learning in Austria's provinces and among different school types. Moreover, some major factors of influence on the level of activity in an e-learning environment are identified in this paper and combined in an explanatory model of usage. According to the results of the binary logistic regression model, the factors province, school type, region and stakeholder of a school as much as the school's total number of users, are major influence factors of the activity at the online platform.

The results of our analysis about topographical and educational differences in the usage behaviour are on the one hand of great value for publishers of the course book and producers of learning software by providing interesting information about the product's distribution in Austria. The results indicate where a stronger promotion of the product is useful and which target groups should be approached. On the other hand, the results are of great relevance for further research steps because they provide an important basis of decision-making for the focus areas and sample selection of further studies in the field of Learning Analytics. The aspect of school type could for example be considered when choosing schools for qualitative interviews or observations in order to figure out why differences in the usage behaviour of different school types exist.

In the second part of this paper, the exercise results of a selected sample were discussed in order to enable decisions on matters such as efficiency and effectiveness of different exercises. An analysis of the average success rate revealed that a large amount of exercises leads to very high average success rates. This paper also contains interesting information about the learning behaviour of students in an e-learning environment. It became apparent that students use to repeat an exercise on the online platform in case their result in the first attempt was low. A repetition usually leads to an improvement of exercise results and increased success rates at the last attempt. Finally, the examination of success rates for the individual exercises of an e-learning course showed that activity types tend to be easier or more difficult as others. The comparison of subjective estimates of difficulty with objective learning results proved to be valuable for finding problems in exercises.

6 Future Work

The results of this study suggest a number of new avenues for research. In general, the authors plan to extend their analysis of exercise results and applied exercises to the population of all schools using the online learning platform more-online.at. This step enables a deeper insight into student's learning behaviour and allows making more generalizing statements. Subsequently further analysis will be performed among the exercises of More! in order to cluster these exercises into homogeneous sets and to allow an extraction of the most valuable exercises. Furthermore, it will be interesting to extend the qualitative analysis of the exercises of More! focusing on technical aspects, usability aspects and teaching methodology. In this regard an inclusion of subject-specific experts in the field of teaching methodology in English education, publishers of the English course book as much as developers of the English learning software will be highly useful. Thereby a better classification of the exercises will be possible allowing more meaningful interpretations of the results attained through this study. Highly valuable exercises could further be used to generate certain profiles of

students in order to enable a better prediction of their level of competence and to enable an easier assessment and a more targeted support and assistance. The research interest thereby focuses on the field of adaptive learning, which aims at an adjustment of computer's presentation of educational material according to certain needs of learners, as indicated by their responses to questions and tasks and is gaining more and more importance in educational technology.

References

[Backhaus et al. 2011] Backhaus, K., Erichson, B., Plinke, W., Weiber, R.: "Multivariate Analysemethoden: Eine anwendungsorientierte Einführung", Springer, Heidelberg, 2011.

[Ben Naim et al. 2009] Ben Naim, D, Marcus N. and Bain, M.: "A User Driven and Data Driven Approach for Supporting Teachers in Reflection and Adaptation of Adaptive Tutorials", in Proceedings of Educational Data Mining 2009: 2nd International Conference on Educational Data, in Cordoba, Mining, Barnes, T., Desmarais, M., Romero, C., & Ventura, S. (Eds.), Cordoba, 21-30, 2009.

[Bitkom 2011] Bundesverband Informationswirtschaft (BITKOM): "Schule 2.0, Eine repräsentative Untersuchung zum Einsatz elektronischer Medien an Schulen aus Lehrersicht", Berlin, 2011. http://www.bitkom.org/files/documents/BITKOM_Studie _Jugend_2.0.pdf (last accessed April 2014)

[BMUKK 2013] Bundesministerium für Unterricht, Kunst und Kultur: "Die Neue Mittelschule", Wien, 2013.
http://www.neuemittelschule.at/themen/kurzinformation_zur_nms.html (last accessed April 2014)

[BMUKK 2014] Bundesministerium für Unterricht, Kunst und Kultur: "Das österreichische Bildungssystem / The Austrian education system", Wien, 2014.
http://www.bmukk.gv.at/medienpool/24501/bildungswege_mai14.pdf (last accessed May 2014)

[Brewster and Fager 2000] Brewster, C., Fager, F., "Increasing Student Engagement and Motivation: From Time on Task to Homework." Portland 2000.
http://home.comcast.net/~reasoned/4410/CRM%20Concept%20Map%20with%20Links/html-tdm-model-hyperlinke_files/motivationforstudents_13.pdf (last accessed May 2014)

[CEFR 2011] Council of Europe: Common European Framework of Reference for Languages: Learning, Teaching, Assessment,
http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf (last accessed May 2014)

[Gerngross et al. 2008] Gerngross, G., Puchta, H., Holzmann, C., Stranks, J., Lewis-Jones, P., Finnie, R.: "MORE! 1 Cyber Homework". Innsbruck, 2008.

[Grasedieck 2008] Grasedieck, D.: "Welchen Wert haben Hausaufgaben? Plädoyer für individualisierte Hausaufgaben"; Neue Didaktik, 1, 25-40, 2008.

[Judd and Kennedy 2004] Judd, T. S., Kennedy, G. E.: "More sense from audit trails: Exploratory sequential data analysis", In Atkinson R., McBeath C., Jonas-Dwyer D., Phillips R. (Eds), Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference, Perth, 5-8 December, 476-484, 2004.

[Kakoty et al. 2011] Kakoty, S., Lal, M., Sarama, S.: "E-learning as a Research Area: An Analytical Approach"; International Journal of Advanced Computer Science and Applications, 2, 9, 144-148, 2011.

[Karakos and Kazanidis 2011] Karakos, A., Kazanidis, I.: "E-learning Platform Usage Analysis" Interdisciplinary Journal of E-learning and Learning Objects, 7, 186-204, 2011.

[Kaur and Krishan 2013] Kaur, M., Kewal, K.: "Cluster Analysis of Behaviour of E-learners", International Journal of Soft Computing and Engineering (IJSCE), Volume 3, Issue 2, 2013.

[Keller and Suzuki 2004] Keller, J. M., Suzuki, K. "Learner motivation and E-learning design: a multinationally validated process", Journal of Educational Media, Vol. 29, No. 3, 229–239, 2004.

[Medienpädagogischer Forschungsverbund Südwest 2013] Medienpädagogischer Forschungsverbund Südwest: "KIM Studie 2012, Kinder + Medien, Computer + Internet, Basisuntersuchung zum Medienumgang 6- bis 13-Jähriger in Deutschland", Stuttgart, 2013. http://www.mpfs.de/fileadmin/KIM-pdf12/KIM_2012.pdf (last accessed April 2014)

[Meier 2006] Meier, R.: "Praxis E-learning, Grundlagen, Didaktik, Rahmenanalyse, Medienauswahl, Qualifizierungskonzept, Betreuungskonzept, Einführungsstrategie, Erfolgssicherung", Gabal, Offenbach, 2006.

[Möller and Wild 2009] Möller, J., Wild, E.: "Pädagogische Psychologie", Springer, Heidelberg, 2009.

[Nandorf and Schmidt 2003] Nandorf, K., Schmidt, T.: "Vom Selbstlernprogramm zum handlungsorientierten, kommunikativen Fremdsprachenunterricht – Überlegungen zum Einsatz lehrwerksbegleitender Software am Beispiel von English Coach 2000 im Englischunterricht", In Legutke, M., Rösler, D. (Eds), Fremdsprachenlernen mit digitalen Medien, Gunter Narr, Tübingen, 2003.

[Neuhold 2013] Neuhold, B.: "Learning Analytics- Mathematik Lernen neu gedacht", Ebner, M., Schön, S. (Eds), Publikationsreihe Internet-Technologie und Gesellschaft, Book on Demand GmbH , Graz, 2013.

[Nomass 2013] Nomass, B.: "The impact of Using Technology in Teaching English as a Second Language"; English Language and Literature Studies, 3, 1, Canadian Center of Science and Education, 111-116, 2013.

[OECD 2005] Organisation for economic co-operation and development (OECD).: "Are Students Ready for a Technology-Rich World? What PISA Studies Tell Us", Paris, 2005. http://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/35995145.pdf (last accessed April 2014)

[Phillips et al. 2011] Philipps, R., Maor, D., Cumming-Potvin, W., Roberts, P., Herrington, J.: "Learning Analytics and study behaviour: A pilot study" In Williams G., Statham P., Brown N., Cleland B. (Eds), Changing Demands, Changing Directions, Proceedings ascilite Hobart, 997-1007, 2011.

[Phillips et al. 2012] Philipps, R., Maor, D., Preston, G., Cumming-Potvin, W.: "Exploring Learning Analytics as Indicators of Study Behaviour"; World Conference on Educational Multimedia, Hypermedia and Telecommunications, Association of Computing in Education (AACE), Colorado, 2861-2867, 2012.

[Sharma 2009] Sharma, P.: "Controversies in using technology in language teaching", 2009. http://www.teachingenglish.org.uk/article/controversies-using-technology-language-teaching (last accessed April 2014).

[Shen et al. 2002] Shen, R., Yang, F., Han, P.: "Data Analysis Center Based on E-Learning Platform", in The internet challenge: Techology and application, Hommel, G. Huanye S. (Eds.), Kluwer Academic Publishers , Dordrecht, 19-28, 2002.

[Shingal 1997] Singhal, M.: "The Internet and Foreign Language Education: Benefits and Challenges", The Internet TESL Journal, 3, 6, 1997. http://iteslj.org/Articles/Singhal-Internet.html (last accessed April 2014)

[Statistik Austria 2013] Statistik Austria: "Bildungsstatistik 2012/13", Wien, 2013. https://www.statistik.at/web_de/statistiken/bildung_und_kultur/formales_bildungswesen/schulen_schulbesuch/index.html (last accessed April 2014)

[Statistik Austria 2011] Statistik Austria: "IKT-Einsatz in Haushalten, Einsatz von Informations- und Kommunikationstechnologien in Haushalten 2011", Wien, 2012. http://www.statistik.at/web_de/dynamic/services/publikationen/17/publdetail?id=17&listid=17&detail=499 (last accessed May 2014)

[Waxman et al. 2003] Meng-Fen, L., Michko Georgette M., Waxmann, H. C.: "A Meta-Analysis of the Effectiveness of Teaching and Learning with Technology on Student Outcomes", North Central Regional Educational Laboratory, 2003. http://treeves.coe.uga.edu/edit6900/metaanalysisNCREL.pdf (last accessed May 2014)